

Modified Flux Balance Analysis

Wei Dai, and Jeffrey D. Varner*

School of Chemical and Biomolecular Engineering

Cornell University, Ithaca NY 14853

Running Title: Modified Flux Balance Analysis

To be submitted: *Biotechnology Journal*

*Corresponding author:

Jeffrey D. Varner,

Professor, School of Chemical and Biomolecular Engineering,

244 Olin Hall, Cornell University, Ithaca NY, 14853

Email: jdv27@cornell.edu

Phone: (607) 255 - 4258

Fax: (607) 255 - 9166

Abstract

The use of computational tools and mathematical modeling has long contributed to our understanding of biochemical networks, in particular, there is a long history of quantitative mechanistic modeling. Currently, there are many existing mathematical approaches to characterize biochemical networks such as cellular metabolism and its regulation. However, many of these methods require detailed kinetic and concentration information that are difficult or even impossible to obtain. In the post-genomics era, large-scale stoichiometric reconstructions of microbial metabolism popularized by static, constraint-based modeling techniques such as flux balance analysis (FBA) have become standard tools. However, FBA models assume a pseudo-steady-state (QSS) approximation and cannot account for the intracellular dynamics of the network. Towards those unmet needs, we present a versatile next-generation flux balance analysis model that is capable of capturing intracellular dynamics and concentrations of the network. The access to intracellular information created allowed us to implement preexisting regulatory methods that could not have been done with previous FBA models. The key innovation of our approach is the integration of kinetic-like constraints that work in combination with the objective function to overcome the issues with unidentifiable parameters or mechanisms to predict flux distributions. We tested our approach by modeling dynamic time evolution of two proof-of-concept networks and an *E. coli* central metabolism network. We were able to replicate classically expected gene expression and enzyme kinetic behavior, including diauxic growth phenomenon. While only small-scale networks with simplified gene expression were evaluated, the framework presented here could be an important first step towards modeling large biochemical networks with undetermined mechanisms or parameters, and complex regulation.

Keywords: Biochemical engineering, systems biology, flux balance analysis

Introduction

The use of computational tools and mathematical modeling has long contributed to our understanding of biochemical networks, in particular, there is a long history of quantitative mechanistic modeling [1]. Currently, there are many existing mathematical approaches to characterize biochemical networks such as cellular metabolism and its regulation [2–11]. However, many of these methods require detailed kinetic and concentration information that are difficult or even impossible to obtain. Even with the growing data bases for cellular components and development of new reduced order modeling approaches [12], the applications of many mechanistic modeling methods are limited by the lack of kinetic and concentration information.

To overcome the lack of kinetic information, alternative course-grained modeling approaches have been used to study flux distributions of metabolic networks. In the post-genomics era, large-scale stoichiometric reconstructions of microbial metabolism popularized by static, constraint-based modeling techniques such as flux balance analysis (FBA) have become standard tools [13]. Since the first genome-scale reconstruction of *Escherichia coli* MG1655 by Edwards and Palsson [14], the reconstruction of over 100 organisms followed [15]. These organisms included prokaryotes such as *E. coli* [16] or *B. subtilis* [17] were highly sought after in bioprocessing to maximize yields of desired products. More recently, network reconstructions have been completed on humans as well [18, 19]. FBA is a course-grained model that relies on a pseudo-steady-state assumption to reduce unidentifiable kinetic models into an underdetermined linear algebraic system that can be solved efficiently even for large systems. Traditionally, FBA models have lack descriptions of metabolic regulation and control mechanisms that cells need for adapting to different chemical and environmental stimulus. The FBA models instead chooses pathways in the network by prescribing an objective function on metabolism. The use of an objective function is crucial for FBA models since predictions are highly dependent on the

objective function used for the analysis. Common objective functions include maximizing growth [13, 20–23], ATP production [24], and the production of desired products [25]. The utilization of objective function in FBA models is a huge advantage over mechanistic models because with little to no kinetic parameters, the maximization of biomass production allows for a wide range of predictions that are consistent with experimental observations for microbial systems [13, 20, 21, 26–30].

In the last decade, FBA have increased in accuracy and utility by incorporating additional biological knowledge from regulatory constraints, thermodynamics, and alternative classes of objective functions [31]. The second generation FBAs initially approached regulatory constraints as Boolean logic operators [29] that arise based on environmental cues. Palsson and coworkers took a bioinformatics approach and incorporated transcription and translation through the use of the E-Matrix to regulate the metabolism [32]. Meanwhile other second generation FBA models focused on incorporating constraints at a physiological level through the use of crowding constraints, enzyme solubility, or membrane economics that constrain the upper limit on the sum of certain or all the flux vectors to increase accuracy of metabolic models [33–35]. The accuracy of the prediction from constraint-based models such as FBA is based heavily on the constraints implemented to reduced the solution space. Second generation FBAs introduce additional constraints that reshape and reduce the solution space to redirect the flux distribution guided by the objective function in a way that more accurately describe biological behaviors.

In this study, we present an effective biochemical network modeling framework for the construction of a next-generation flux balance analysis. The key innovation of our approach is the integration of kinetic-like constraints that work in combination with the objective function to overcome the issues with unidentifiable parameters or mechanisms to predict flux distributions. These kinetic-like constraints are adjustable parameters that are governed by the presence any combination of: substrate, enzyme, and signaling

molecules. This integration allows the description of fluxes and complex regulatory interactions such as allosteric regulation of enzyme activity in the absence of specific kinetic and mechanistic information which limited traditional FBA models. We adopted the regulatory rules and method from [11], which do not rely on overarching theoretical abstractions or restriction assumptions. We tested our approach by modeling the time-evolution of a hypothetical proof-of-concept metabolic network using discretized mass balances. In particular, we verified that our new modeling approach can replicate classically expected gene expression and enzyme kinetic behavior, and capture intracellular metabolite concentrations and regulatory behaviors. Towards these questions, we constructed a proof-of-concept metabolic network that has two enzyme-dependent reactions. In the absence of the specific enzymes, the reactions will not occur, a behaviors that were that were not captured by the traditional FBA models. We found that our next-generation FBA method has the potential to incorporate gene expression, complex regulatory interactions, and capture intracellular metabolite concentration. Lastly, we tested this framework with a central carbon network for *E. coli* and was able to adjust minimal parameters to fit experimental data of [36] and capture the diauxic effect of carbon utilization. However, with minimal experimental measurements and known parameters, this modeling approach fails to capture intracellular metabolites at an adequate level. However, we believe that this framework presented here could be an important addition to the next-generation FBA models.

Results

1. Formulation of Proof-of-concept Network Model We developed a proof-of-concept metabolic network (Fig. 1) to investigate the features of our effective biochemical network modeling approach. This ultra-reduced network incorporates metabolic reactions, gene expression, enzyme-dependent reactions, and allosteric regulation. The two examples presented results in the consumption and utilization of an extracellular substrate \mathcal{A}_{xt} in a series of metabolic reactions. Several of these reactions or fluxes involved are enzyme or initiator-dependent. The fluxes that were not explicitly enzyme-dependent assumed enzyme saturation and that the flux is only a function of substrate concentration.

The conversion from \mathcal{A} to \mathcal{B} , and the extracellular uptake of \mathcal{C}_{xt} is dependent on the presence of *enzyme* \mathcal{E} , and \mathcal{C} transporter (\mathcal{CT}), respectively. The expression of \mathcal{E} and \mathcal{CT} is regulated at the gene expression level. Gene 1 is transcribed and translated into *enzyme* \mathcal{E} in the presence of growth factors, while Gene 2 transcription and translation results in \mathcal{CT} in a similar manner at a lower concentration of \mathcal{A}_{xt} . In the presence of a growth factor that acts as an inducer, the gene for *enzyme* \mathcal{E} will activate and facilitate transcription and translation reactions to synthesize *enzyme* \mathcal{E} . Meanwhile, in the presence of \mathcal{A}_{xt} , the gene expression of \mathcal{CT} is repressed. Once the extracellular \mathcal{A} is depleted, the \mathcal{C} transporter will be synthesized. In our formulation, Hill-like transfer functions were used to calculate the influence of factor abundance on target enzyme activity. In this context, factors can be a function of metabolite concentration or flux. The kinetic-like constraints are multiplied by an activity control between 0 and 1.

Gene expression in the proof-of-concept network involved is a two-step process: transcription and translation. The coarse-grained transcription and translation reactions in our network neglected the utilization of NTPs and amino acids. Transcription in the proof-of-concept network is solely determined by the presence of the relevant inducer, respective gene, and RNA polymerase concentration, and translation reactions were determined by

mRNA and ribosome concentration. We assumed the generation of RNA polymerase and ribosome synthesis to be a zeroth order constant while degradation to be a first order rate.

A critical test of our modeling approach was to simulate networks with known behavior. If we cannot reproduce the expected behavior of simple networks, then our effective modeling strategy and particular rules of constraining fluxes will not be feasible for large-scale networks. We considered two cases that had identical initial concentrations and kinetic parameters but with different objective functions and gene expression. The first case involved the synthesis of \mathcal{B}_{xt} from \mathcal{A}_{xt} , a case analogous to the maximization of a desired product while minimizing any undesired byproducts. The second case we tested included the addition of a biomass template reaction and we explored the diauxic growth phenomenon through a simplified representation of carbon catabolite repression.

1.1 Dynamics of inducers and enzyme dependent reactions The following study focused on the impact of growth factor (inducer) on networks aimed at the production of \mathcal{B}_{xt} . Metabolic Gene 2 (Fig. 1) responsible for the expression of \mathcal{CT} was not included. The growth factor concentration was varied systematically in a linear fashion and the dynamic changes in system was determined. We implemented the objective function to maximize the production of \mathcal{B}_{xt} . The intracellular metabolites concentrations were constrained with a lower limit of zero while the flux bounds were implemented as a function of substrate concentration, control parameters, and enzyme concentration for the reactions that were enzyme dependent. The biomass was kept at a constant of 0.1 gDW/L without any growth. Gene expression of *enzyme* \mathcal{E} was governed by the concentration of the inducer, a growth factor. With the increased inducer concentrations, we observed an expected increase in mRNA and enzyme concentration until steady state (Fig. 2A).

The presence of Enzyme E facilitated the conversion of \mathcal{A} to \mathcal{B} . The rate of conversion was governed by the concentration of both of the Enzyme E and species \mathcal{A} . The rate of conversion of \mathcal{A} to \mathcal{B} is shown in Fig. 2B. At an inducer concentration of zero, the rate

of conversion is zero (not shown). Initially, the rate of \mathcal{A}_{xt} uptake was greater than the consumption of species \mathcal{A} (Fig. 2C). As *enzyme* \mathcal{E} accumulated, the conversion of \mathcal{A} to \mathcal{B} increased proportionally to *enzyme* \mathcal{E} 's and substrate \mathcal{A} 's concentration. The total rate of consumption \mathcal{A} was greater than the generation, causing it to deplete until intracellular \mathcal{A} reaches steady state.

The gene expression of *enzyme* \mathcal{E} redirected the pathway of synthesis for product \mathcal{B}_{xt} . In the absence of *enzyme* \mathcal{E} , \mathcal{B} can only be formed from the longer path of conversion of \mathcal{A} to \mathcal{C} to \mathcal{B} (Fig. 1). This pathway leads to an increased production of undesired byproduct \mathcal{C}_{xt} (Fig. 2D). The presence of Enzyme E opens a new pathway for the direct synthesis of \mathcal{B}_{xt} through the conversion of \mathcal{A} to \mathcal{B} (Fig. 1) and reduces the accumulation of intracellular \mathcal{C} and byproduct \mathcal{C}_{xt} (Fig. 2D).

1.2 Dynamics of Growth and Alternative Substrate Utilization The proof-of-concept network was used to test whether our approach is capable of capturing the diauxic growth phenomenon observed in bacteria using gene expression to regulate the metabolism. A biomass template reaction of intracellular metabolite consumption was constructed:



The growth factor that is responsible for the expression of *enzyme* \mathcal{E} which catalyzes the conversion of \mathcal{A} to \mathcal{B} was maintained constant at zero. The objective function was implemented to maximize biomass production and the species and flux constraints were implemented the same way as the previous case.

We used an initial biomass concentration of 0.1 gDW/L, and a growth rate constraint set between to initiate the simulation with 10 arbitrary units of \mathcal{A}_{xt} . The extracellular product, \mathcal{C}_{xt} is generated as a byproduct from the production of all the biomass precursors (Fig. 3A) since the presence of species \mathcal{C} drives the flux of \mathcal{C} to \mathcal{C}_{xt} . As \mathcal{A}_{xt} is depleted,

the repression of \mathcal{CT} gene expression decreases gradually in a sinusoidal behavior with respect to \mathcal{A}_{xt} concentration. When all of the \mathcal{A}_{xt} was consumed, gene expression of \mathcal{CT} became uninhibited and the production of \mathcal{CT} allowed the network to uptake \mathcal{C}_{xt} and utilize it as an alternate substrate for growth (Fig. 3 A and B).

With the proof-of-concept model, we were able to capture the diauxic growth phenomenon with regulation at the level of gene expression. We witnessed a decrease in intracellular species concentration due to the decreasing substrate uptake rate relative to the metabolic demand of growth. However, when \mathcal{A}_{xt} was depleted, intracellular \mathcal{B} and \mathcal{C} began to accumulate due to the relatively high \mathcal{C}_{xt} uptake rate compared to that of \mathcal{A}_{xt} uptake rate before complete depletion. As \mathcal{C}_{xt} was consumed, the intracellular concentrations decreased again (Fig. 3D).

In addition, we witnessed that the intracellular concentration of \mathcal{B} depleted rapidly due to biomass demand (Fig. 3D). Species \mathcal{B} is synthesized by the conversion from \mathcal{C} , and is consumed by the conversion of \mathcal{B} to \mathcal{A} and biomass synthesis, and diluted due to growth. Intracellular \mathcal{B} is the last species in the linear pathway (Fig. 3E), the mFBA is optimized to only provide enough \mathcal{B} for growth and dilution which kept the \mathcal{B} concentration at minimal. Once \mathcal{A}_{xt} is depleted, \mathcal{B} accumulated and facilitated the conversion of \mathcal{B} to \mathcal{A} to provide the biomass demand for \mathcal{A} . \mathcal{A} is the last species in the linear pathway, similarly, \mathcal{A} concentration became minimal. Our method was able to reduce the solution space, successfully capturing complex trends by using simple Hill function and kinetic-like constraints.

1.3 Initialization of E. coli Diauxic model As stated before, a critical test of our modeling approach was to simulate networks with known behavior. We decided to implement our framework on the central carbon metabolism of E. coli. In addition, we wanted to test the potential of our modeling approach when the kinetic parameters, mechanistic equations, and regulatory effects within the network architecture were not known. We formed

simple regulatory rules based on observation of the experimental data [36] of diauxic growth: the uptake of acetate was inhibited at high levels of glucose uptake. Our *E. coli* model, consist of 65 species and 111 reactions including the template biomass reaction adopted from [37] (Fig. 4A). For the initialization of intracellular metabolite concentration, we used measured values from different literature sources [37, 38]. For the kinetic-like flux constraints, we decided to implement the classic Michaelis Menten (MM) kinetics to capture any saturation effects within the system since the metabolite concentration are well above saturation in vivo [37].

A critical challenge for any dynamic ODE model is the estimation of kinetic parameters. For metabolic processes, there is also the added challenge of identifying the regulation and control structures that manage metabolism. This problem still exists with our FBA modeling approach, but this challenge can be overcome by increasing the uncertainty of our kinetic-like flux constraints. By doing so, the distance between the upper and lower flux increased and therefore increase the solution space and the optimized solution becomes more heavily influenced by the objective function. In our system, we decided to take a course-grained approach for parameter estimation to test the capacity and adaptability of our modeling framework. Using our course-grained parameter estimation approach, the Michaelis Menten parameter were evaluated based on the initial condition of the cellular system. A standard FBA was then carried out, optimizing for growth rate. The calculated solution to the fluxes was used to evaluate V_{max} using initial substrate concentrations and assuming a K_m of 1/50 of the species' initial concentration. For pathways that were not utilized, ie flux of zero, in the standard FBA solution, the flux constraints were unbound due to the uncertainty. To account for the error and uncertainty associated with this crude estimation, an uncertainty tolerance was placed between the lower and upper bound as a constraint that the flux value must fall within. To determine this variation within the flux constraint, we varied a global parameter across all flux constraints that

were governed by MM to fit experimental data. Most of the fluxes in our system were not sensitive to the change solution space. In most metabolic networks, only a few reactions are rate limiting and act as a bottleneck for the whole system. Increasing the solution space did not impact the fluxes downstream of the bottleneck. This robustness within our system is similar to the findings of Sethna and coworkers who indicate that the model's performance is often controlled by only a few parameter combinations, a characteristic that seems universal for multi-parameter models referred to as sloppiness [39]. Regulatory control was put in place for the uptake of acetate by utilizing a Hill-function governed by the flux of glucose uptake rate of the previous iteration of the discretized mass balance. The parameters in the fill functions were determined by manually fitting against experimental data. The pathways of gluconeogenesis shown in Fig. 4A were not regulated, but rather controlled by the optimization from the objective function.

Under the assumed condition, our model was able to accurately capture experimental measurements of extracellular glucose and acetate, and the biomass and glucose-acetate diauxic phenomenon (Fig. 4B). However due to the crude parameter estimation, some dynamic resolution of the intracellular concentration were lost. Most of the intracellular metabolites either accumulated or depleted to the concentration limits that were set for the intracellular species. On closer examination, the two intracellular byproducts of growth, formate and acetate (Fig. 4A), accumulated towards their respective concentration threshold. However, formate depleted rapidly, reaching its lower threshold concentration after it reached the maximum threshold concentration. Upon the depletion of glucose, the state of the system changed to the utilization of acetate leading to a new set of flux distribution. The flux of the glycolysis and gluconeogenesis are shown in Fig. 4D. The flux of the glycolysis pathway shown remained constant throughout growth with the utilization of glucose, upon glucose depletion, the flux of gluconeogenesis took over in supplying carbon into the upper glycolysis and pentose phosphate pathways for biomass precursor

production.

Discussion

In this study, we presented an effective kinetic-like constraint-based FBA model to dynamically simulate metabolic networks. Our proposed strategy integrated a kinetic modeling approach with an objective function from the traditional FBA to dynamically describe metabolic regulation and control. We tested this model approach by developing a hypothetical proof-of-concept network and a central carbon *E. coli* network. In particular, we tested whether our effective modeling approach could describe classically expected behaviors through the combination of regulatory control from allosteric regulation and gene expression. Towards these questions, we explored two hypothetical and one realistic networks. In each of the two hypothetical networks, substrate \mathcal{A}_{xt} was transported into the cell and consumed for the production of \mathcal{B}_{xt} or biomass depending on the objective function. The *E. coli* network consumed glucose for the production of biomass precursors for growth. The reactions were assumed to be only substrate-dependent unless specifically implemented to be a function of substrate and enzyme concentrations. When these flux constraints were integrated into the network models, these constraints captured dynamic patterns in the flux and concentration in the network. Lastly, we implemented this method on a central carbon metabolic network of *E. coli*, and captured the diauxic growth phenomenon.

This proposed modeling technique shares features with other modeling techniques. The implementation of the flux constraints is flexible and can take the form of most kinetic or mechanistic forms, ie mass action kinetics or effective kinetic formats [40]. This new modeling approach takes the form of a discretized form of effective kinetic model in the way that flux constraints that are implemented. The overall purpose of our framework functioned in a similar manner as other FBA models to reduce the solution space to what is biologically or thermodynamically relevant. The real innovation of this modeling approach is imploring the use of intracellular metabolite. This subtle difference allowed us

to integrate preexisting an effective kinetic modeling approach [11, 12] into our kinetic-like constraints. In addition, having access to intracellular species information allowed us to integrate allosteric regulation into the model.

This proposed modeling framework also differs appreciably from previous established FBA models. The vast difference between our proposed method and an ODE models involves the uncertainty that sets tolerance between the governing kinetic form and the upper and lower bound of the flux constraint and the guidance of the objective function. Even with unidentifiable kinetic parameters, this combination of high uncertainty and objective function of our modeling approach can overcome this problem. Based on the fluxes that were determined, the kinetic parameters of the most optimized flux distribution could be determined and analyzed for feasibility. Unlike the previous FBA models, we did not make the quasi-steady-state (QSS) assumption for intracellular species. Our internal species was allowed to accumulate and deplete at a constant rate each time the discretized mass balances were evaluated so that the intracellular concentrations could be captured and implemented as contributors for allosteric regulation. At high uncertainty of the kinetic parameters, we were still able to effectively capture glucose consumption, growth, and the diauxic phenomenon. Unfortunately, dynamic resolution for intracellular metabolites was lost as most of the intracellular metabolites approached their upper and lower limits. However, with enough certainty in our parameters, we were able to fully capture the dynamics of intracellular metabolite in our proof-of-concept networks.

In summary, our proposed framework offers an effective method of reducing the solution space of the FBA by using kinetic-like approximations that captured simple enzyme dependence, allosteric regulation, and diauxic growth. At its core, our effective modeling approach combines advantages of FBA and kinetic ODE models. The use of an objective function that acts as a global predictor will overcome problems pertaining unidentifiable parameters or mechanisms, meanwhile the use of kinetic-like constraints on the fluxes

acts as a local check-points that reshapes the solution space that is closer to the true behavior of the network. There is a trade off between uncertainty and resolution as expected for any model. At the uncertainty of our system increases, our mFBA approaches the state of a tradition FBA where the solution is determined from the objective function. Meanwhile as uncertainty of our system approaches zero, our modeling approach takes a form that is more similar to an ODE model that is discretized and the solution is shaped by kinetic-like flux constraints. We believe that utilization of this modeling approach is for large networks with unidentifiable parameters or unsalable using ODEs.

Materials and Methods

The Modified Flux Balance Analysis (mFBA) is new method created for the study of biochemical networks much like the traditional flux balance analysis. The mFBA utilizes series of species constraints and flux constraints similar to the traditional FBA. The mass balances for each species are discretized and solved using explicit numerical methods such as explicit Euler as in **Equations** (2), (3), and (4).

$$\frac{dx}{dt} = S * v \quad (2)$$

$$\text{IntracellularSpecies} : x_{t+1} = x_t + b\Delta t - x_t\mu\Delta t \quad (3)$$

$$\text{ExtracellularSpecies} : x_{t+1} = x_t + b\Delta tX \quad (4)$$

In equation 2, S is the stoichiometric matrix with m species by n reactions, v is a flux vector of each reaction, μ denotes the growth rate and b is a vector of change in species with respect to time. In the discretized mass balances, **Equations** (3) and (4), t is the time, μ is the growth rate, x denotes the species concentration, and X is the cellmass. For traditional FBA, the intracellular species (IS) were assumed to be at quasi-steady-state (QSS), making the species constraint, $b = 0$. However, for the mFBA, we assumed an initial concentration for intracellular species concentration based on literature measurements [41,42] and did not assume QSS by allowing intracellular metabolites to accumulate and deplete within a physiological range. By incorporating the lower and upper limit the species concentration and dilution due to growth, the species constraints implemented in our model is shown in **Equation** (5).

$$LB - x(1 - \mathbf{dsa}\mu) \leq S * v \leq UB - x(1 - \mathbf{dsa}\mu) \quad (5)$$

In **Equation 4** and are vectors of lower and upper limits of each species, while represents the dilution selection array, a Boolean vector that takes determines whether a species is diluted due to growth, ie. 0 for extracellular species that are not diluted and 1 for intracellular species that are.

The flux constraints use two key components. The kinetics portion are implemented as a function of species concentration in a generalized fashion, and then the regulation portion is incorporated for regulatory mechanisms. The flux is constrained between two a lower and upper bound (**Equation (6)**).

$$\alpha \leq v \leq \beta \quad (6)$$

For tradition FBA, and are a vector of constants same length as the number of reactions, for our modeling approach, they are functions of control parameters, substrate concentration, and/or enzyme concentration as shown in **Equation (7)**:

$$\alpha = f_1(\epsilon)v \quad \beta = f_2(\epsilon)v_k \quad (7)$$

where and are uncertainty functions of , the uncertainty coefficient, a constant that is used to control the solution space in the model based on the certainty in the parameters in the kinetic-like constraints, . The uncertainty functions act as a way to widen the distance between lower and upper limit defined by kinetic-like constraints. The kinetic-like constraints, can take any form, from mechanistic descriptions to effective kinetic models like the power-law:

$$v_k = \alpha x^n * C \quad (8)$$

where \mathbf{g} and \mathbf{c} represents vectors of gain coefficients and cooperativity coefficients of the power law respectively. The vector \mathbf{u} is a control vector that utilizes Hill-functions to describe any types of regulation within the network for both activation and repression.

$$\text{Activation : } \mathcal{C} = \frac{\beta x^n}{1 + \beta x^n} \quad \text{Repression : } \mathcal{C} = 1 - \frac{\beta x^n}{1 + \beta x^n} \quad (9)$$

Similarly, \mathbf{g} and \mathbf{c} represents vectors of gain coefficients and cooperativity coefficients. Lastly, the network is subjected to the objective function, :

$$Z = \sum_{i=1}^{n_{fluxes}} c * v \quad (10)$$

The vector \mathbf{v} was used to select a linear combination of metabolic fluxes that were included in the objective function. In our examples, we have defined either production of desired species or cellular growth to be our objective function.

Acknowledgements

This study was supported by an award from [FILL ME IN].

References

1. Bailey JE (1998) Mathematical modeling and analysis in biochemical engineering: past accomplishments and future opportunities. *Biotechnology progress* 14: 8–20.
2. Castellanos M, Wilson DB, Shuler ML (2004) A modular minimal cell model: purine and pyrimidine transport and metabolism. *Proceedings of the National Academy of Sciences of the United States of America* 101: 6681–6686.
3. Fell D, Cornish-Bowden A (1997) *Understanding the control of metabolism*, volume 2. Portland press London.
4. Domach M, Leung S, Cahn R, Cocks G, Shuler M (1984) Computer model for glucose-limited growth of a single cell of *Escherichia coli* b/r-a. *Biotechnology and bioengineering* 26: 203–216.
5. Barkal N, Leibler S (1997) Robustness in simple biochemical networks. *Nature* 387: 913–917.
6. Liao JC (1993) Modelling and analysis of metabolic pathways. *Current opinion in biotechnology* 4: 211–216.
7. Novák B, Tóth A, Csikász-Nagy A, Györfy B, Tyson JJ, et al. (1999) Finishing the cell cycle. *Journal of Theoretical Biology* 199: 223–233.
8. Kompala DS, Ramkrishna D, Jansen NB, Tsao GT (1986) Investigation of bacterial growth on mixed substrates: experimental evaluation of cybernetic models. *Biotechnology and Bioengineering* 28: 1044–1055.
9. Palsson B, Joshi A, Ozturk S (1987) Reducing complexity in metabolic networks: making metabolic meshes manageable. In: *Federation proceedings*. volume 46, pp. 2485–2489.
10. Conrado RJ, Mansell TJ, Varner JD, DeLisa MP (2007) Stochastic reaction–diffusion simulation of enzyme compartmentalization reveals improved catalytic efficiency for a synthetic metabolic pathway. *Metabolic engineering* 9: 355–363.

11. Wayman JA, Sagar A, Varner JD (2015) Dynamic modeling of cell-free biochemical networks using effective kinetic models. *Processes* 3: 138–160.
12. Sagar A, Varner JD (2015) Dynamic modeling of the human coagulation cascade using reduced order effective kinetic models. *Processes* 3: 178–203.
13. Lewis NE, Nagarajan H, Palsson BO (2012) Constraining the metabolic genotype–phenotype relationship using a phylogeny of in silico methods. *Nature Reviews Microbiology* 10: 291–305.
14. Edwards J, Palsson B (2000) The escherichia coli mg1655 in silico metabolic genotype: its definition, characteristics, and capabilities. *Proceedings of the National Academy of Sciences* 97: 5528–5533.
15. Feist AM, Herrgård MJ, Thiele I, Reed JL, Palsson BØ (2009) Reconstruction of biochemical networks in microorganisms. *Nature Reviews Microbiology* 7: 129–143.
16. Feist AM, Henry CS, Reed JL, Krummenacker M, Joyce AR, et al. (2007) A genome-scale metabolic reconstruction for escherichia coli k-12 mg1655 that accounts for 1260 orfs and thermodynamic information. *Molecular systems biology* 3.
17. Oh YK, Palsson BO, Park SM, Schilling CH, Mahadevan R (2007) Genome-scale reconstruction of metabolic network in bacillus subtilis based on high-throughput phenotyping and gene essentiality data. *Journal of Biological Chemistry* 282: 28791–28799.
18. Duarte NC, Becker SA, Jamshidi N, Thiele I, Mo ML, et al. (2007) Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proceedings of the National Academy of Sciences* 104: 1777–1782.
19. Thiele I, Swainston N, Fleming RM, Hoppe A, Sahoo S, et al. (2013) A community-driven global reconstruction of human metabolism. *Nature biotechnology* 31: 419–425.
20. Mahadevan R, Edwards JS, Doyle FJ (2002) Dynamic flux balance analysis of diauxic

- growth in *escherichia coli*. *Biophysical journal* 83: 1331–1340.
21. Burgard AP, Maranas CD (2003) Optimization-based framework for inferring and testing hypothesized metabolic objective functions. *Biotechnology and bioengineering* 82: 670–677.
 22. Edwards JS, Ramakrishna R, Palsson BO (2002) Characterizing the metabolic phenotype: a phenotype phase plane analysis. *Biotechnology and bioengineering* 77: 27–36.
 23. Van Dien SJ, Lidstrom ME (2002) Stoichiometric model for evaluating the metabolic capabilities of the facultative methylotroph *methylobacterium extorquens* am1, with application to reconstruction of c3 and c4 metabolism. *Biotechnology and bioengineering* 78: 296–312.
 24. Ramakrishna R, Edwards JS, McCulloch A, Palsson BO (2001) Flux-balance analysis of mitochondrial energy metabolism: consequences of systemic stoichiometric constraints. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 280: R695–R704.
 25. Varma A, Boesch BW, Palsson BO (1993) Biochemical production capabilities of *escherichia coli*. *Biotechnology and bioengineering* 42: 59–73.
 26. Schilling CH, Covert MW, Famili I, Church GM, Edwards JS, et al. (2002) Genome-scale metabolic model of *helicobacter pylori* 26695. *Journal of bacteriology* 184: 4582–4593.
 27. Edwards JS, Ibarra RU, Palsson BO (2001) In silico predictions of *escherichia coli* metabolic capabilities are consistent with experimental data. *Nature biotechnology* 19: 125–130.
 28. Ibarra RU, Edwards JS, Palsson BO (2002) *Escherichia coli* k-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature* 420: 186–189.
 29. Covert MW, Palsson BØ (2002) Transcriptional regulation in constraints-based

- metabolic models of escherichia coli. *Journal of Biological Chemistry* 277: 28058–28064.
30. Edwards JS, Palsson BO (2000) Robustness analysis of the escherichiacoli metabolic network. *Biotechnology Progress* 16: 927–939.
 31. Kauffman KJ, Prakash P, Edwards JS (2003) Advances in flux balance analysis. *Current opinion in biotechnology* 14: 491–496.
 32. Thiele I, Jamshidi N, Fleming RM, Palsson BØ (2009) Genome-scale reconstruction of escherichia coli's transcriptional and translational machinery: a knowledge base, its mathematical formulation, and its functional characterization. *PLoS Comput Biol* 5: e1000312.
 33. Beg QK, Vazquez A, Ernst J, de Menezes MA, Bar-Joseph Z, et al. (2007) Intracellular crowding defines the mode and sequence of substrate uptake by escherichia coli and constrains its metabolic activity. *Proceedings of the National Academy of Sciences* 104: 12663–12668.
 34. Shlomi T, Benyamini T, Gottlieb E, Sharan R, Ruppin E (2011) Genome-scale metabolic modeling elucidates the role of proliferative adaptation in causing the warburg effect. *PLoS Comput Biol* 7: e1002018.
 35. Zhuang K, Vemuri GN, Mahadevan R (2011) Economics of membrane occupancy and respiration-fermentation. *Molecular systems biology* 7: 500.
 36. Varma A, Palsson BO (1994) Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type escherichia coli w3110. *Applied and environmental microbiology* 60: 3724–3731.
 37. Bennett BD, Kimball EH, Gao M, Osterhout R, Van Dien SJ, et al. (2009) Absolute metabolite concentrations and implied enzyme active site occupancy in escherichia coli. *Nature chemical biology* 5: 593–599.
 38. Heijnen J (2010) Impact of thermodynamic principles in systems biology. In: *Biosys-*

tems Engineering II, Springer. pp. 139–162.

39. Machta BB, Chachra R, Transtrum MK, Sethna JP (2013) Parameter space compression underlies emergent theories and predictive models. *Science* 342: 604–607.
40. Heijnen JJ (2005) Approximative kinetic formats used in metabolic network modeling. *Biotechnology and bioengineering* 91: 534–545.

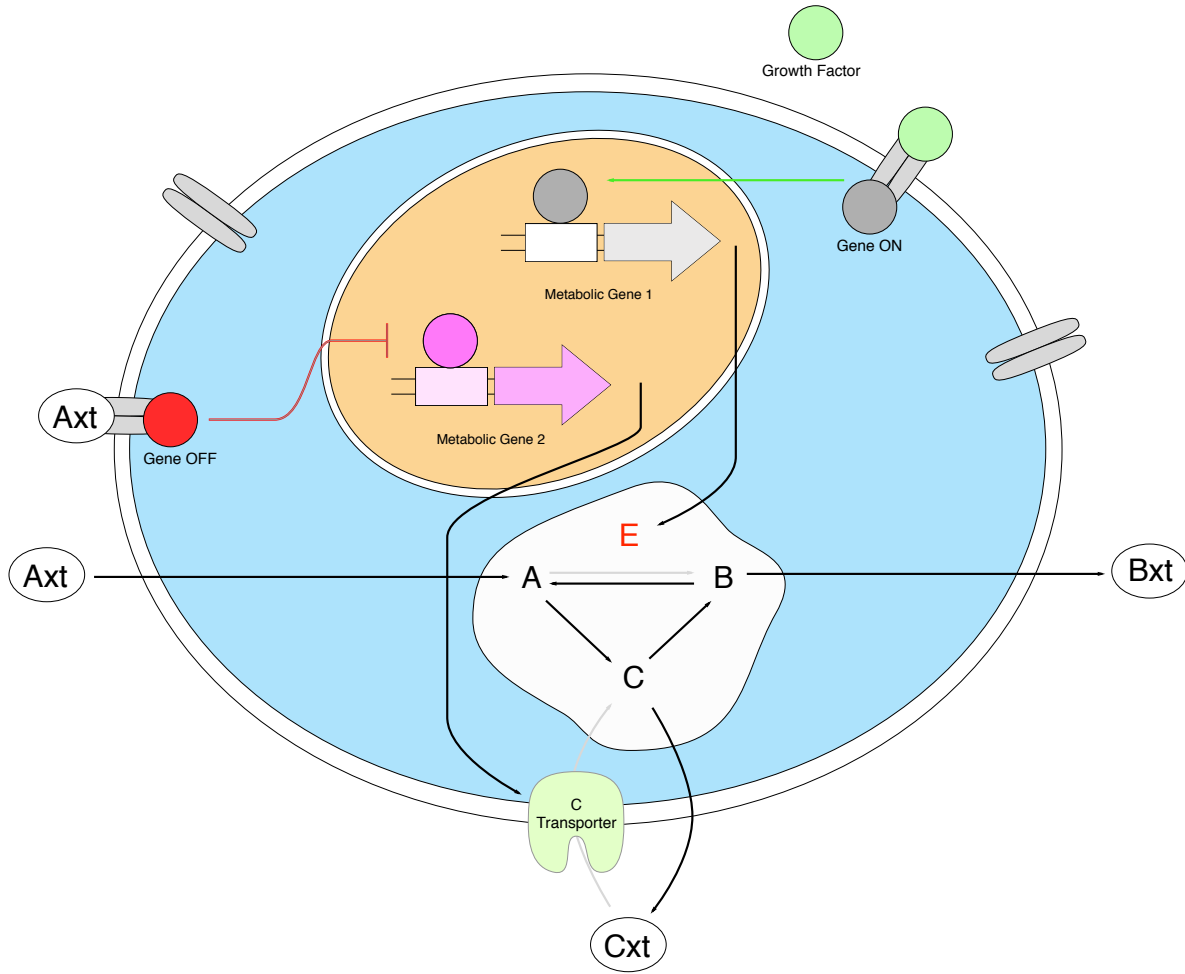


Fig. 1: Proof-of-concept metabolic network considered in this study. The black arrows indicate depletion and formation, the red arrows indicate repression, the green represents activation, and the grey arrows are enzyme-dependent reactions. Substrate A_{xt} can be transported into the cell and converted into B and C to be transported out. The reactions that are enzyme dependent are subjected to regulation with respect to gene expression (*enzyme \mathcal{E}* and *Ctransporter*).

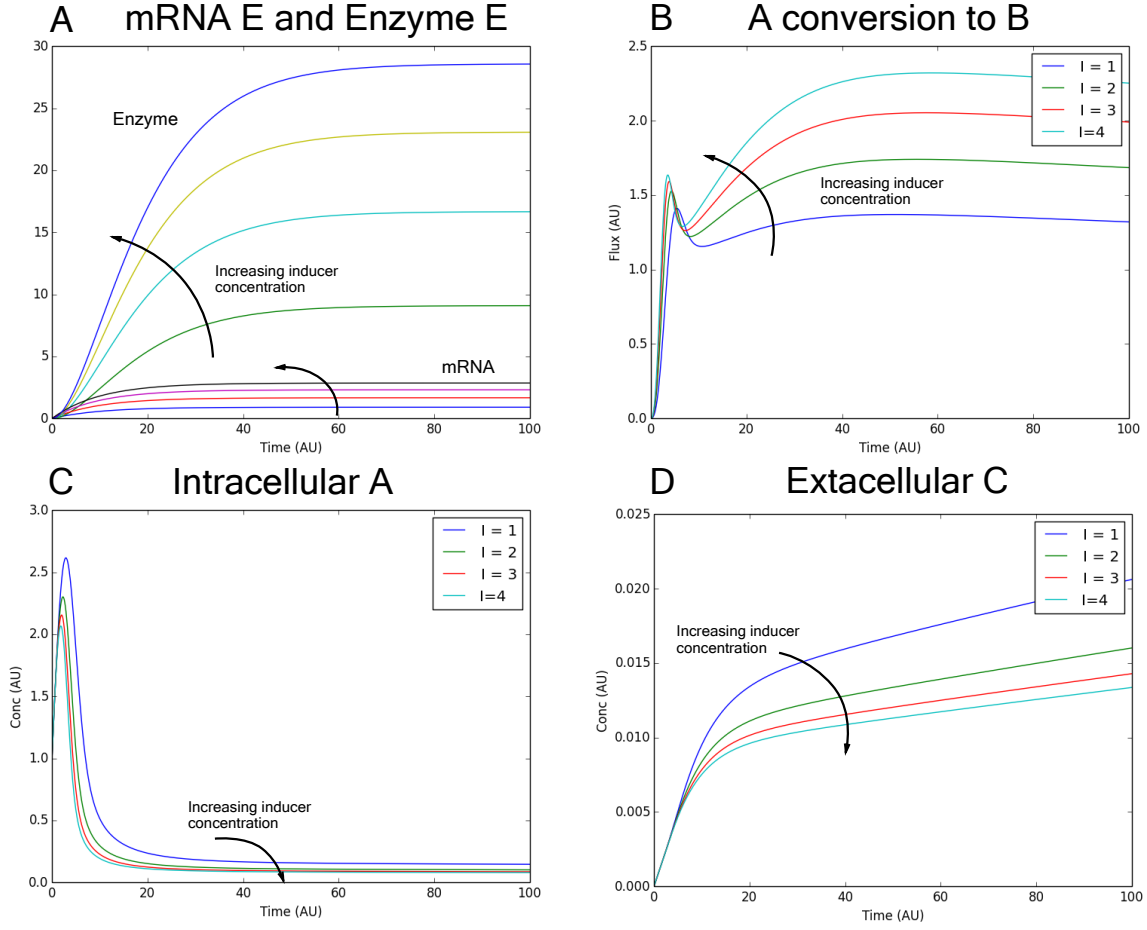


Fig. 2: (A): mRNA E and Enzyme E response for different growth factor concentrations. Inducer concentration of zero did not express any mRNA or Enzyme (not shown). **(B):** Flux of A to B catalyzed by *EnzymeE*. The rate of A conversion to B is proportional to the Enzyme E and substrate A concentration. **(C):** Intracellular concentration of A. **(D):** C_{xt} concentration in the absence of *Ctransporter*. The production of C_{xt} decreases with increasing growth factor inducer concentration.

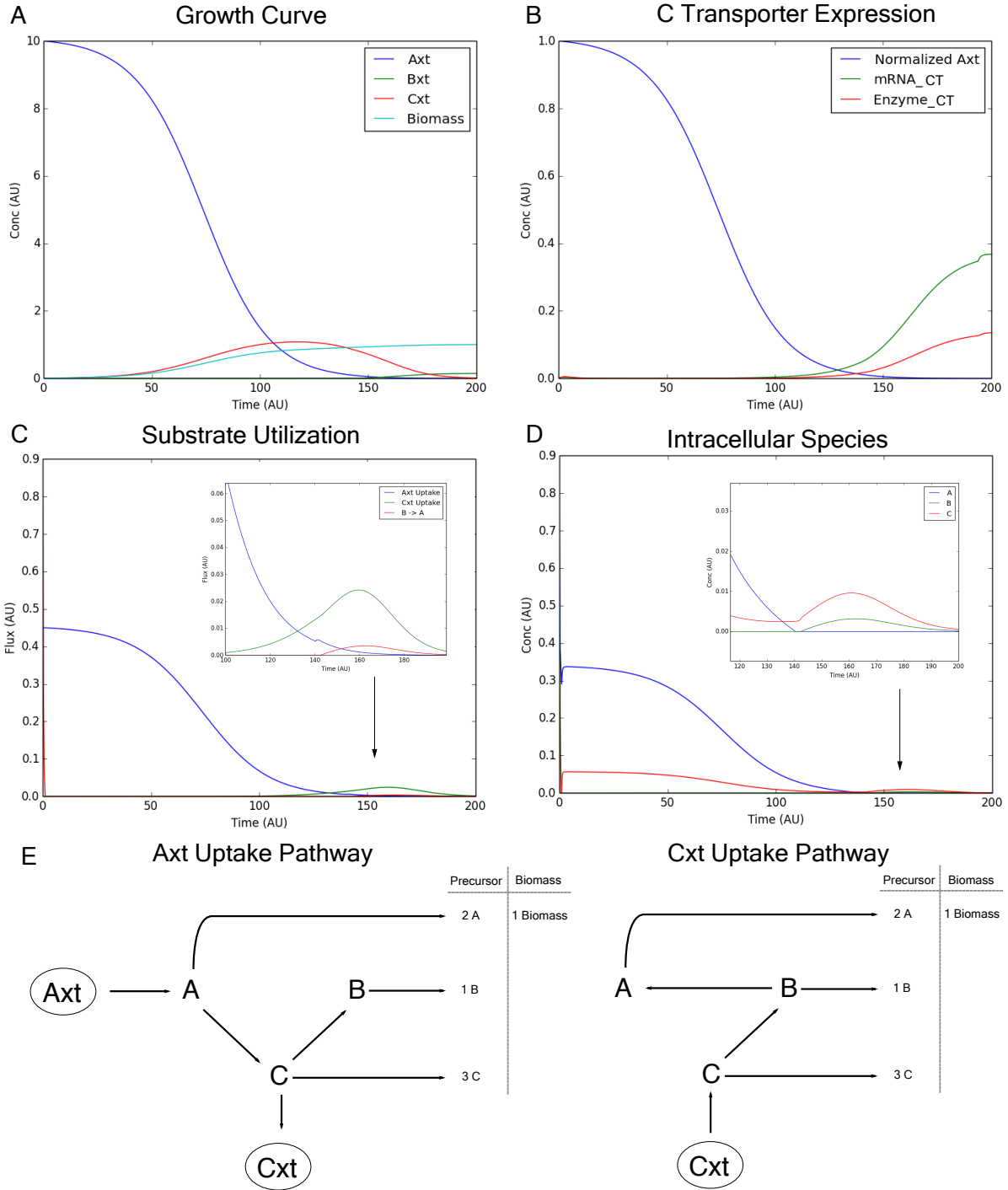


Fig. 3: (A): Growth Curve indicates the utilization of C_{xt} as an alternate substrate when A_{xt} is depleted. **(B)** Production of mRNA *Ctransporter* and *Ctransporter* enzyme begins once the A_{xt} concentration is below its inhibition threshold concentration. **(C):** The conversion of B to A is required for intracellular replenishment for growth upon A_{xt} depletion. **(D):** Concentration of intracellular species. **(E):** Linear pathway of metabolic flux. While A_{xt} is available, intracellular A , C and B are first, second, and third species respectively in the linear pathway. Once A_{xt} is depleted and C_{xt} is utilized as an alternate substrate, the order of the pathway changes to C , B , and A .

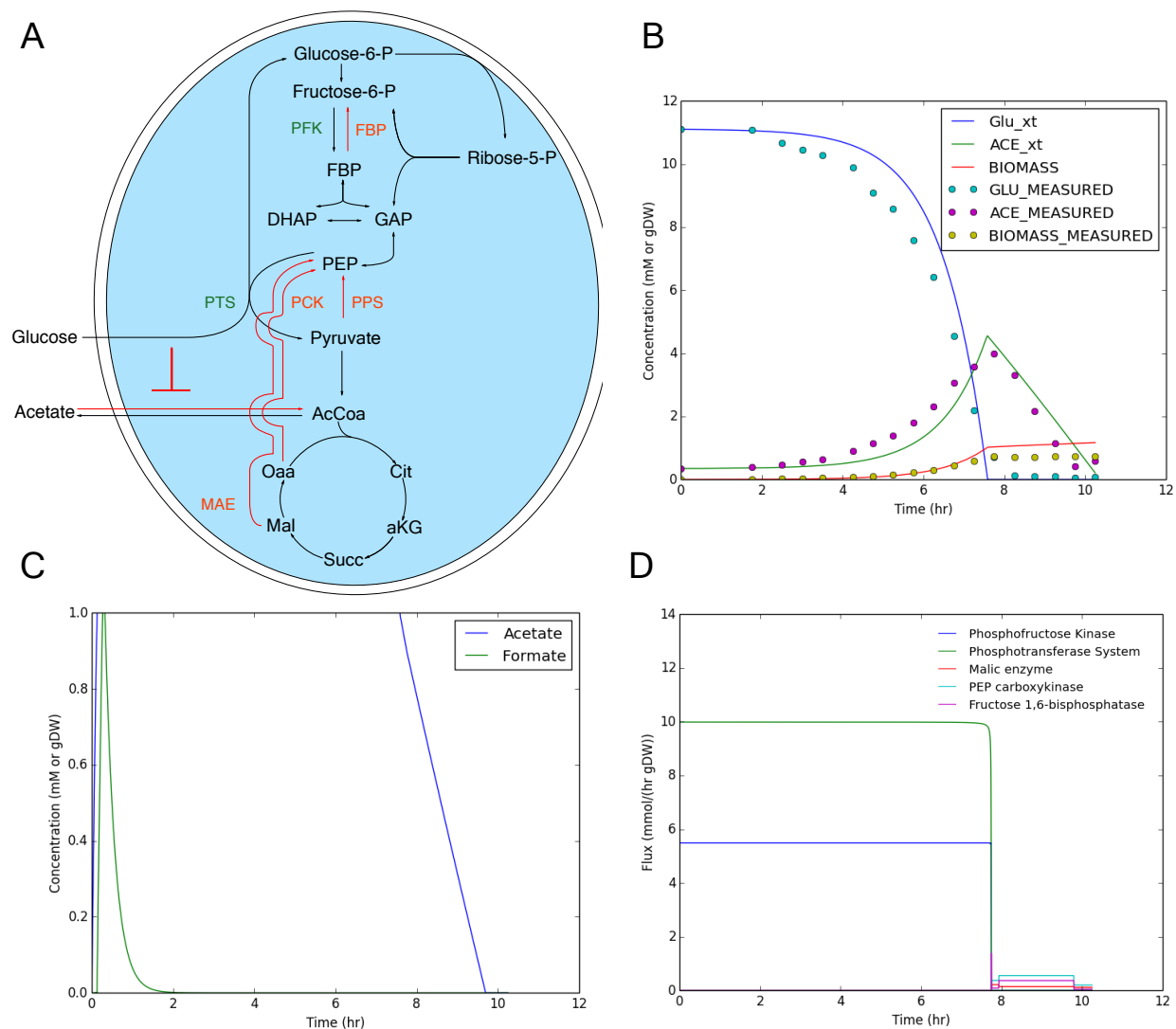


Fig. 4: (A): Simplified illustration of the E.coli central metabolism. The red arrows represent the fluxes for gluconeogenesis. (B): Growth curve showing utilization of acetate compared against experimental measurements from Varma et al [REF]. (C): Intracellular metabolite concentration of two by products, acetate and formate. (D): Flux of glycolysis and gluconeogenesis.